# Drivers of high-latitude plant diversity hotspots and their congruence

Niskanen, Annina Kaisa Johanna

2017-08-01

Niskanen , A K J , Heikkinen , R K , Väre , H U & Luoto , M 2017 , ' Drivers of high-latitude plant diversity hotspots and their congruence ' , Biological Conservation , vol. 212, Part A , pp. 288-299 . https://doi.org/10.1016/j.biocon.2017.06.019

http://hdl.handle.net/10138/307297
https://doi.org/10.1016/j.biocon.2017.06.019

# Drivers of high-latitude plant diversity hotspots and their congruence

Annina K.J. Niskanen[1], Risto K. Heikkinen[2], Henry Väre[3], Miska Luoto[1]

Corresponding author: Annina K.J. Niskanen, e-mail: annina.k.niskanen@helsinki.fi, phone:+358504635706

[1] Department of Geosciences and Geography, P.O. Box 64 (Gustaf Hällströmin katu 2a), FI-00014 University of Helsinki, Finland; [2] Finnish Environment Institute, Natural Environment Centre, P.O. Box 140, FI-00251, Helsinki, Finland; [3] Finnish Museum of Natural History, Botanical Museum, University of Helsinki, P.O. Box 7, Room 131.1, Unioninkatu 44, Helsinki, Finland

**Abstract**

Determining the drivers, patterns and hotspots of biodiversity can be of critical importance in supporting regional conservation planning. However, as biodiversity hotspots can be described with several different metrics, it is important to investigate their congruence as well as the spatial overlap of hotspots with protected areas. Here, by using extensive data on climate, topography, soil characteristics and vascular plants combined with boosted regression tree models, we determine the patterns and key drivers of plant diversity hotspots along broad environmental gradients in northernmost Europe spanning from taiga landscapes to treeless tundra. We assess plant diversity with four metrics – species richness, range-rarity richness, threatened species richness, and local contribution to beta diversity – and examine their congruence with each other as well as with contemporary conservation areas.

We found that climate plays an important role in governing species diversity, though topo-edaphic are highlighted alongside climatic predictors in determining the diversity patterns of many threatened, near-threatened, and range-restricted species. Importantly, the different diversity metrics have contrasting drivers and, overall, their hotspots have low congruence. Furthermore, existing protected areas appear to offer limited coverage for hotspots of vascular plant diversity.

Modelling the various facets of diversity and their drivers, such as the topo-edaphic setting, may help conserve diversity in a changing climate. Projected patterns of different aspects of diversity and their congruency can provide insights for the processes underlying biodiversity and be employed to assess the representativeness of protected area networks.

Keywords: biodiversity; hotspot; conservation; species richness; LCBD

## 1. Introduction

The successful conservation of biodiversity depends on our ability to understand and predict the properties and distribution of diversity and, in particular, the hotspots thereof. This calls for continuous efforts to determine the key drivers of biodiversity (Gould 2000, Moritz 2001, Zellweger et al. 2015). Identifying biodiverse regions, i.e. diversity hotspots, can also be useful in recognizing priority areas for conservation (Prendergast et al. 1993, Myers et al. 2000). Hotspots can be identified via assessments of diversity patterns across landscapes. They are commonly quantified through different diversity metrics, such as measures of the total, rare, narrow-ranged, or threatened species richness occurring at a given site (Prendergast et al. 1993, Williams et al. 1996, Reid 1998, Armsworth et al. 2004, Ceballos and Ehrlich 2006). Total species richness, combined with rarity or endemism data, has traditionally had a central role in conservation assessments (Myers et al. 2000, Stein et al. 2000), but the degree to which diversity features overlap remains contradictory (Bonn et al. 2002, Orme et al. 2005). A less investigated but potentially valuable diversity metric is the measure of a site's local contribution to β-diversity (LCBD: Legendre and Cáceres 2013, Legendre 2014), which has yet to be applied in terrestrial plant diversity hotspot analyses. LCBD highlights ecologically unique sites contributing more than the mean to β-diversity. An important application in mapping

different diversity hotspots is the comparison of their distributions with existing Protected Areas (PAs) to determine potential conservation shortfalls (Scott et al. 1993, Flather et al. 1997, Virkkala et al. 2013, Huang et al. 2016).

Total species richness (TSR), a direct proxy for α-diversity (i.e. local diversity at a given site), is widely used to depict the biodiversity and conservation value of an area (Whittaker 1972, Steck et al. 2007, Magurran and Dornelas 2010). However, TSR may be insufficient in representing local aggregations of rare species (Reid 1998). This is a potentially critical deficiency as spatially restricted species and species of elevated conservation concern contribute greatly to biological uniqueness and are often classified as having a greater extinction risk than common species (Csuti et al. 1997, Lamoreux et al. 2006, Peters et al. 2015). Thus, hotspots of such species are considered to have a higher conservation value than areas where species richness is similar but composed of mainly common species (Lennon et al. 2004, Malcolm et al. 2006, Mouillot et al. 2013). The richness of spatially restricted species can be represented by range-rarity richness (RRR), also known as rarity-weighted richness and rarity score. RRR is a frequently used biodiversity metric to quantify and locate areas richest in species with the most restricted ranges (Williams et al. 1996, Myers et al. 2000, Levin et al. 2007). The richness of species of high conservation concern, including both the threatened and near threatened species (THR) of a given site can be exemplified by the amount of Red Listed species present (Gjerde et al. 2004). TSR, RRR and THR are commonly used in diversity studies (Bonn et al. 2002, Orme et al. 2005), but often utilized separately. A more recent metric, LCBD indicates a site's relative contribution to overall β-diversity (the variation in species composition across sites: Anderson et al. 2011). A high LCBD may indicate distinctive conditions or combinations of species with high conservation value, or degraded species-poor sites in need of ecological restoration (Legendre and Cáceres 2013, Legendre 2014).

Identifying the drivers of biodiversity and assessing differences between diversity indicators can help focus field surveys and conservation efforts, or aid in detecting threats to biodiversity (Brooks et al. 2006, Cañadas et al. 2014). Previous studies have shown that various environmental factors influence species richness patterns (Wohlgemuth 1998, Lobo et al. 2001, Loidi et al. 2015) and that observed biodiversity hotspots generally showcase low congruence (e.g. Feng et al. 2011, Daru et al. 2015). There is thus growing interest in finding a more comprehensive way to identify diversity hotspots, but knowledge is still lacking regarding what drives different diversity metrics, their hotspots and congruence, and how these manifest across extensive environmental gradients at high latitudes (Orme et al. 2005, Magurran 2013).

Here we address these information gaps by seeking further understanding on which factors govern vascular plant diversity patterns and the congruence of different diversity hotspots in high-latitude continental Europe. This study combines a statistical modelling approach with extensive data of regional environmental attributes and a unique dataset of vascular plant species based on field observations. We examine the drivers and patterns of four diversity metrics: overall species richness (Prendergast et al. 1993), range-rarity richness (Myers et al. 2000), richness of species of elevated conservation concern (Gjerde et al.

2004), and local contribution to β-diversity (Legendre and Cáceres 2013). By determining the effects of climatic, topographic and edaphic parameters along a broad gradient spanning from forested taiga to treeless tundra, we identify what drives discernable patterns and differences between the metrics. Furthermore, we quantify the congruence of diversity hotspots to examine their value for conservation efforts. Lastly, as hotspots falling outside areas currently protected can help define conservation gaps in a geographically explicit way (Flather et al. 1997), we overlay our results with current PAs to evaluate their efficiency in preserving vascular plant diversity within the studied high-latitude region.

## 2. Materials and methods

### *2.1. Study area*

The study area, located in northernmost Europe between 67°N and 69°N, is influenced by the Arctic Ocean, the proximal Scandes Mountains (Fig. 1), the Polar Front, and the warm North Atlantic current (Aalto et al. 2014). The region hosts a variety of climatic, topographic and edaphic gradients and has an elevational gradient of 72 to 1365 m.a.s.l.. Average July temperatures range from 6.1 °C to 15.2 °C and mean annual precipitation from 449 mm to 600 mm (1981 – 2010 means: Pirinen et al. 2012). The study area encompasses a boundary area between northern boreal and arctic-alpine habitats, and the vegetation varies accordingly from spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) forests in the south to mountain birch (*Betula pubescens* subsp. *czerepanovii*) and alpine tundra above the tree-line in the north (Sormunen et al. 2011, le Roux et al. 2012). See Niskanen et al. (2016a, 2016b) for further details on the study area.

### *2.2. Vascular plant species data*

Occurrence records for 593 vascular plant species in 1 × 1 km cells ($n$ =2073) served as the basis for this study (Fig. 1). Species data was collected for each cell by professional botanists and supplemented by exploration of species records from scientific literature and herbaria with the floristic material being maintained in the Kastikka-database, property of the Botanical Museum (University of Helsinki, Finnish Museum of Natural History). Ranging from taiga to treeless tundra, all the main biotopes found in the study region are represented by the species data. The northern parts of the study region were subjected to a higher sampling intensity. To account for the possible effects of this sampling bias, we utilized spThin, an R package for spatially thinning species occurrence records for use in SDMs (Aiello-Lammens et al. 2015). The data were thinned 100 times with points randomly removed from within a given radius (here, 5 km). The resulting thinned data samples retained 214 sites each.

Based on the species data, we calculated four easily replicated diversity metrics (TSR, RRR, THR and LCBD: Table 1) to capture different aspects of biodiversity. Metric equations and corresponding details are listed in Table 1 and the summary statistics are presented in Table 2. The TSR metric was calculated as the total number of vascular plant species occurring in a given grid cell (Prendergast et al. 1993, Gaston 2000). The RRR metric, following terminology suggested by Guerin and Lowe (2015), was built up by first calculating the range-rarity richness of a single species as the inverse of range size (Williams 2000), here

substituted with the estimated range cover of each species within Finland, Sweden, and Norway according to the Nordic Flora (Mossberg and Stenberg 2003) . The range-rarity richness of a 1 × 1 km cell was calculated as the sum of the inverse of the ranges of all the species occurring in that cell (Table 1). Due to a strong correlation with TSR, we divided the range-rarity richness of each cell by its TSR to give relative range-rarity richness (RRR: suggested by Williams et al. 1996). RRR thus employs grid cell occupancy to quantify range restriction, with higher values indicating greater rarity (Williams et al. 1996, Levin et al. 2007).

The THR metric was based on the Red List status of the species (Rassi 2010). The categories employed here are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), and Near Threatened (NT). A total of 73 species (12%) in our data set fall within these categories: three as CR, ten as EN, 23 as VU, and 37 as NT (Appendix A). THR thus represents the total number of the species of elevated conservation concern occurring in a given cell. The LCBD metric was calculated according to the methodology proposed by Legendre and Cáceres (2013). We utilized the R scripts provided therein, adjusted with Sørensen dissimilarity coefficients and tested with 999 permutations. Large LCBD values indicate highly unique sites in terms of community composition, i.e. they contribute more than the mean to β-diversity (see Legendre and Cáceres 2013 for further details).

*2.3. Environmental data*

We utilized an extensive environmental data set of climatic, topographic, and edaphic variables resampled to a 1 × 1 km resolution covering the entire study region (n= 25 766) to identify dominant diversity predictors (Table 2). Important ecophysiological conditions relating to temperature, water, light, and nutrients are generally considered as vital for plants (Guisan and Zimmermann 2000, Austin and Van Niel 2011b, Mod et al. 2016). Thus, we compiled a set of environmental variables presumed to affect vascular plant diversity (Scherrer and Körner 2011, Reside et al. 2014) which we then tested for pair-wise correlations using the Spearman's rank correlation test (Appendix B). Variables with correlation coefficients lower than |0.7| were considered for further analysis (Dormann et al. 2013) and nine variables were finally selected to represent climate, topography, and soil properties (Table 2).

The climatic variables, obtained from national observation networks in Finland, Norway and Sweden (Finnish Meteorological Institute; The Norwegian Meteorological Institute 2012; Swedish Meteorological and Hydrological Institute 2012, respectively), are based on the normal period of 1981–2010, representing current climatic conditions. Monthly mean temperature and precipitation values were modeled using generalized additive models incorporating geographical position, water cover, and topography at a 200 × 200 m resolution (Aalto and Luoto 2014). Here, we included three climatic variables: (1) freezing degree days (FDD; average overwintering conditions); (2) growing degree days (GDD3; average growing conditions); and; (3) water balance (WAB; range of available moisture). FDD and GDD3 are based on the effective temperature sum of mean daily temperatures below 0 °C or above 3 °C, respectively (Carter et al. 1991,

Fronzek et al. 2006). WAB was determined as the difference between the mean annual precipitation sum and potential evaporation (Skov and Svenning 2004).

Topography can indirectly influence plants by altering growing conditions (Austin and Van Niel 2011a, Moeslund et al. 2013). Here, the topographic variables are based on an Aster -derived digital elevation model (DEM: NASA Land Processes Distributed Active Archive Center (2013); spatial resolution 30 × 30 m). Three topographic variables were selected: (1) range of slope angle representing slope processes and topographical variability; (2) incoming potential solar radiation representing surface temperature conditions (McCune and Keon 2002); and, (3) topographic wetness index (TWI), a proxy for soil moisture accounting for upslope moisture availability (Beven and Kirkby 1979). Slope angle was calculated with the Spatial analyst extension in ArcGis 10.0. Radiation was calculated with ArcView 3.2 Solar analyst extension accounting for latitude, elevation, slope angle and aspect, daily and seasonal shifts in solar angle, atmospheric attenuation, and topographical shadows (Fu and Rich 1999). TWI was calculated with a Python script that accounts for slope and upslope contributing area, written by Prasad Pathak (Beven and Kirkby 1979, Esri 2013). These variables are common proxies for rugged terrain microclimates (Guisan and Zimmermann 2000, Dobrowski 2011), soil moisture (Penna et al. 2009), or geomorphological processes (Randin et al. 2009) that can influence high-latitude vegetation properties (see e.g. le Roux et al. 2013a, le Roux et al. 2013b).

Edaphic parameters influence vegetation through soil properties (Guisan et al. 1998; Austin and Van Niel 2010). Three edaphic variables were used in this study: (1) calcareousness, i.e. proportion of nutrient-rich bedrock representing soil pH, shown to improve species distribution model predictive power (Dubuis et al. 2013); (2) substrate evenness to represent growing substrate heterogeneity (rock, sand, peat and/or till); and, (3) rock cover (cliffs, rocky outcrops, scree), which may be critical for modelling species in severe environments (Guisan et al. 1998). The edaphic variables were reclassified from a digital database (Geological Survey of Finland 2010; spatial resolution 20 × 20 m) and transformed following Aalto and Luoto (2014). To calculate evenness, we first calculated Simpson's D, a common diversity index:

$$\left(D = \frac{1}{\sum {p_i}^2}\right)$$

where $p_i$ is the proportion of substrate types relative to the total number of substrate types. From this we calculated a common measure of evenness and diversity, Simpson's E:

$$\left(E = \frac{D}{S}\right)$$

where $D$ is Simpson's D and $S$ is the number of soil types present in a given grid cell.

181 *2.5. Modelling environmental drivers of diversity and identifying hotspots*

182 We applied boosted regression tree (BRT) modelling to the four diversity metrics to identify
183 their drivers and distributions (Fig. 2). BRT modelling combines the strengths of boosting and multiple
184 regression trees. BRT models can fit complex nonlinear relationships, do not require prior data transformation
185 or outlier elimination, and take into account interactions between predictors. BRT models have a high
186 predictive performance and can be summarized to provide deep ecological insight (Elith et al. 2008).

187 In the first phase of the modelling process, relationships between observations (100 separate
188 thinned samples of 214 cells each) of the four diversity metrics and the nine explanatory variables were fitted
189 using functions from the gbm package (Ridgeway 2010). Each BRT model was built by setting tree complexity
190 to four, learning rate to 0.005 and the minimum number of observations in terminal nodes to two. The number
191 of trees was limited to 3000, and bagging fraction to 0.5. The response variables were fitted with identical sets
192 of predictors using Poisson distributions. We used a metric of $R^2$ as suggested by Kissling and Carl (2008) to
193 assess the fit of the model, hereafter called $R^2$. The importance, i.e. relative influence; sensu Elith et al. (2008),
194 of a predictor was assessed from a combination of the frequency the variable was selected as a model predictor
195 and the improvement resulting from the inclusion of the variable (Elith et al. 2008).

196 In the second modelling phase, we produced predictive maps for the four diversity metrics by
197 fitting the models to cover all grid cells in the study region (n = 25766). The models were calibrated using all
198 grid cells with available species data within a data sample. Models were cross-validated based on four runs,
199 each time selecting a 70% random calibration data sample while verifying model accuracy against the
200 remaining 30%. We assessed model predictive power by comparing the observed and predicted values of the
201 evaluation data using $R^2$ analysis.

202 In the third phase, we applied model predictions to delineate diversity hotspots. Hotspots were
203 subjectively defined as the richest 5% of cells (n = 1289 for each metric), a commonly used threshold (see e.g.
204 Prendergast et al. 1993, Reid 1998, Myers et al. 2000). Overall and pairwise hotspot congruence are analyzed
205 by measuring the extent of spatial overlap (Orme et al. 2005). Pairwise hotspot congruence is also measured
206 by calculating the true skill statistic (TSS; Allouche et al. 2006) and Cohen's kappa (Cohen 1960) between the
207 different diversity hotspots using R statistical software and the PresenceAbsence package. TSS and kappa and
208 range from -1 to +1, with values $\leq 0$ demonstrating the agreement between metrics to be equivalent to or worse
209 than chance, and values close to 1 representing perfect agreement. As the TSS and kappa scores gave similar
210 results, we only present TSS values here.

211 Finally, to assess the congruence between hotspots and PAs, we employed data on terrestrial
212 conservation areas from the World Database on Protected Areas (Chape et al. 2005, UNEP-WCMC 2016) and
213 overlaid this with our predictions (Scott et al. 1993, Huang et al. 2016). All statistical analyses were carried
214 out in R version 3.2.2 (R Development Core Team 2015). See Figure 2 for a modelling overview.

## 3. Results

The mean number of observed vascular plant species per cell was 60 species (67 in the thinned data), ranging from six to 240 (219) species. Relative range-rarity richness ranged from near zero to 6.2 (3.5) with a mean of 0.3 (0.03). The average richness of species of elevated conservation concern was less than two species per cell, and varied between zero and 32 (14) species. Values of LCBD ranged from 0.0003 to 0.0007, averaging out at 0.0005 (see Table 2).

The combined effect of the four most important drivers accounted for a majority of variable relative influence. TSR showed a particularly strong relationship with climate, as GDD3 for a majority of mean variable importance (60%: Fig. 3). TSR hotspots are most likely to occur in areas with warmer growing conditions. Thus although TSR hotspots exhibit a scattered distribution throughout the study region, a significant portion is found in the south (Fig. 4). For RRR, GDD3 is the most influential variable (53% mean variable importance), followed by TWI, radiation and FDD (combined mean relative influence of 26%). RRR hotspots are most likely to occur in areas with, on average, cooler growing seasons and milder winters, combined with a high TWI (Fig. 3). They are highly clustered and concentrated almost entirely in the northernmost parts of the study region (Fig. 4). THR is, besides the main influence from GDD3, also strongly affected by topo-edaphic drivers (Fig. 3), with hotspots predicted to be found in areas with calcareous substrates and high TWI and WAB (Fig. 4). LCBD is mainly influenced by two climatic predictors, GDD3 and FDD (Fig. 3). The spatial pattern of LCBD is patchy, with hotspots located mainly in the northern regions (Fig. 4).

The mean fit of the models, i.e. model explanatory power as demonstrated by $R^2$, were 0.65 (σ 0.02) for TSR, 0.66 (σ 0.13) for RRR, 0.57 (σ 0.03) for THR, and 0.45 (σ 0.19) for LCBD (Fig. 3). Model predictive ability (cross-validated $R^2$ values), were 0.46 (σ 0.08) for TSR, 0.30 (σ 0.18) for RRR, 0.23 (σ 0.07) for THR, and 0.04 (σ 0.05) for LCBD (Fig. 2; see Appendix C for the model predictive ability boxplot). These results suggest that the more commonly used richness metrics can be explained fairly well with our suite of predictors, whereas LCBD is more challenging to model. However, the model outcomes display limited predictive accuracy, particularly for TSR, THR and LCBD.

Our hotspot analysis reveals little spatial congruence between the metrics (Figs 4 & 5). Spatial congruence is moderate between TSR and THR (TSS = 0.52): sites with high species richness are more likely to host species of elevated conservation concern than the other metrics. Some congruence was found between RRR and LCBD (TSS = 0.11). Most of the hotspots are discordant (TSS = -0.05) and do not share any hotspots (Figs 4 & 5). Cumulatively, hotspots for the four metrics occupied a total of 4255 grid cells (representing ~17% of the study area: Fig. 4). Overall hotspot congruence is extremely low: 4% were congruent for two metrics; none of the hotspots overlapped for more than two metrics (Figs 4 & 5).

Current terrestrial PAs cover 38% of the study region. (Fig. 6). The metrics most congruent with PAs are RRR and LCBD, with 1039 and 1037 hotspot grid cell predictions falling within current PAs,

respectively. TSR and THR have 151 and 179 of their hotspots within PAs, respectively (Fig. 6). In total, 2149 hotspots fall into PAs. Thus, 50% of the predicted hotspots are currently protected. However, hotspot congruence with existing conservation areas varied markedly for the different diversity metrics. Of all the hotspots falling within contemporary PAs, 257 are congruent for two indices, representing 12% of the congruent hotspots.

## 4. Discussion

Our results show that large variations in vascular plant diversity characterize high-latitude landscapes. Climate plays an important role in regulating these patterns, but the spatial configuration of diversity-rich locations is critically dependent on the metric used to assess diversity. In addition, topo-edaphic variables appear important in determining patterns of species of high conservation concern or those with limited range size. From a conservation planning perspective it is important to acknowledge that the predicted diversity hotspots are generally highly non-congruent. Moreover, although our study region has an extensive PA network, it provides only varying coverage and appears highly limited for hotspots of total species richness and species of elevated conservation concern.

### *4.1. The patterns and drivers of diversity*

Our findings demonstrate that parallel to climate and energy availability generating diversity gradients at global scales (Jetz and Rahbek 2002, Hawkins et al. 2003), patterns of high-latitude plant diversity are mainly climatically delineated even at finer scales (Fig. 3). Energy availability during the growing season appears particularly instrumental for overall richness patterns, predominantly realized through widely distributed species (Davies et al. 2007), as well as the richness of species with restricted ranges. However, the diversity metrics show different responses to climate: contrary to TSR, RRR exhibits an adverse response to a milder climate. Climate is usually a determinant of large-scale species patterns, whereas topo-edaphic factors influence smaller scale distributions (Pearson and Dawson 2003): our results suggest that patterns of TSR are more influenced by having suitable conditions satisfied at higher hierarchical levels (i.e. climate) with only minor effect from further fulfilment of finer scale habitat requirements. As inspections of hotspot locations and drivers could enable diversity threat detection (Cañadas et al. 2014), it follows that due to the strong effect by climate on all of the predicted hotspots, these hotspots may be particularly sensitive to climatic change. Though a useful predictor for TSR, RRR and LCBD, climate appears slightly less influential for predicting patterns of threatened or near-threatened plant species (see Jetz and Rahbek 2002).

THR, though also affected by climate, appears to also be reliant on conditions further down the hierarchy, exhibited by topo-edaphic speciation through the influence of slopes and substrate (Fig. 3). Our findings show calcareous bedrock to support a higher THR (see also Heikkinen and Neuvonen 1997, Anderson and Ferree 2010). As calcareous areas are not common in this region, it follows that many of the species favoring a high soil pH are rare or threatened (Kauhanen 2013). Furthermore, due to linkages between bedrock pH and plant diversity with evolutionary history (Chytrý et al. 2003), historical dynamics are highly relevant,

alongside contemporary conditions, for patterns of plant diversity (Gaston 2000). Areas with higher soil moisture are also more likely to be THR hotspots by providing suitable habitats for greater numbers of at-risk species. Moister soils can reduce exposure to extremes in temperature (Ashcroft and Gollan 2013) and decouple local climates from broad-scale conditions (le Roux et al. 2013a, Lenoir et al. 2016), thus slowing projected temperature changes (Ackerly et al. 2010). Topo-edaphic conditions may thus buffer species from unfavorable changes, remaining constant relative to dynamic climatic factors (Reside et al. 2013). This could have implications particularly for the species of high conservation concern, predicted to bear the brunt of climate change (Bomhard et al. 2005, Thuiller et al. 2005). In other words, the more topo-edaphically defined THR hotspots may be more likely to have a higher chance of being buffered from upcoming changes than the more climatically influenced hotspots. Furthermore, in a warming climate, the hotspots offering cooler temperatures (RRR) may be more adversely affected than those with a positive response to temperatures (TSR and THR).

Although we find that diversity patterns can be modelled with environmental variables, similarly to previous studies at high latitudes (Heikkinen and Neuvonen 1997, Niskanen et al. 2016a), predictions have limited surrogacy for observations. The limited predictive accuracy indicates problems requiring further attention in modelling future diversity patterns, particularly of species of high conservation concern or local contributions to diversity. The low to moderate predictive power could be partly explained by the divergent environmental requirements of the large set of common and rare species used here (Heikkinen et al. 1998, Jetz and Rahbek 2002). The deterministic selection of species assemblages by specific environmental factors can be seen particularly in the environmental variables of the highest importance (here, mainly climatic drivers, local topography, and substrate). However, the low accuracy of some of the models suggests that some stochastic or ecologically neutral processes acting here are insufficiently captured with our modelling framework. Furthermore, the effects of disturbance regimes (which can affect biodiversity patterns, see le Roux and Luoto 2014) were not accounted for here due to the lack of robust spatial disturbance data covering the whole study area . Nevertheless, we acknowledge that such processes are likely to be relevant for future studies of cold-climate biodiversity. Overall, our results support the importance of direct observations for biodiversity assessments and highlight a need for further knowledge regarding the complex interactions between, not only species and the environment, but also various diversity metrics.

*4.2 The congruence of diversity hotspots*

The different drivers behind the various aspects of diversity are mirrored in the discordant spatial patterns and low hotspot congruence (Figs 4 & 5; see also Orme et al. 2005). The low congruence between the metrics could result from contrasting responses to environmental conditions, further intensified by how TSR hotspots are more likely to share species with other sites and thus have lower LCBD (see e.g. Maloufi et al. 2016). Limited overall hotspot congruence could also reflect the small size of the hotspots, or the chosen hotspot threshold (Daru et al. 2015). Low congruence leads to difficulties in sustaining various facets of diversity simultaneously, indicating that multiple measures are necessary to capture the complexity

of biodiversity and that an integrative approach is needed in identifying priority areas. For example, though they have the highest total species count, using TSR hotspots as priority sites would potentially exclude the most ecologically unique sites and hotspots of range-restricted species (Fig. 4). Consequently, the different hotspots have contrasting utility for conservation efforts as hotspot surrogacy cannot be assumed (see Possingham and Wilson 2005, Feng et al. 2011, Daru et al. 2015). The low overall congruence further emphasizes the importance of the congruent hotspots for conservation (Figs & 5). The partial congruence of TSR and THR hotspots shows how, in some cases, the conservation of overall species diversity might inadvertently be beneficial for threatened or near-threatened species. However, knowledge of the limited congruence between metrics should not be neglected in future studies.

*4.3 Diversity hotspots and protected areas*

Regardless of substantial PA coverage, our results illustrate how looking at spatial extent alone gives a rather simplistic indication of area effectiveness in conserving biodiversity: only a fifth of the predicted hotspots are currently protected (Fig. 6). There is particularly little association between hotspots of TSR and THR with PAs, suggesting that the current reserve network may not be sufficient for safeguarding the regions' overall plant diversity or species of elevated conservation concern (Fig. 6). On the other hand, RRR and LCBD hotspots show a higher rate of congruence with existing PAs. This congruence, possibly resulting from an inclination of these hotspots to occur in the more northern, cooler climes, may, however, be challenged due to the warming climate. The contribution of RRR and LCBD hotspots to biological and ecological uniqueness (Csuti et al. 1997, Legendre 2014, Peters et al. 2015) and their congruence suggest potential complementary utility for conservation. Along with most of the protected hotspots not being congruent, only 257 (29%) of the hotspots where different diversity aspects coincide are found in PAs. These results, along with the knowledge that only a small proportion of these PAs are strictly protected (most were established in accordance to the Wilderness Act), emphasize a need for further evaluations of the effectiveness of currently protected areas. As many conservation frameworks prioritize overlapping areas of diversity (Brooks et al. 2006), simply using congruent hotspot locations as future conservation targets could be a robust but simple conservation option. However, as contemporary hotspots are not necessarily future hotspots (Moritz 2001), additional analyses of overlapping hotspots, with a more dynamic approach in regards to utilizing knowledge on their drivers or how they will be affected by climate change, is warranted.

One approach outlined in a review of climate change adaptation strategies to conserve biodiversity (Mawdsley et al. 2009) is the identification and protection of refugia, i.e. sites buffered from climatic changes aiding biodiversity persistence (Taberlet and Cheddadi 2002, Barnosky 2008, Ashcroft 2010). Linkages have already been found between hotspots of endemism (Harrison and Noss 2017) and phylogeography (Médail and Diadema 2009) with past climate refugia. Hotspot and refugia research both call for an increased understanding of connections between biodiversity and the environment (Schut et al. 2014), and combining knowledge of high-latitude refugia (Niskanen et al. 2016a) with the hotspot approach could be valuable for biodiversity conservation (sensu Médail and Diadema 2009).

## 5. Conclusions

The drivers and differences between diversity metrics are highly relevant, interlinked areas of research that could lead to a more effective balance between ecosystem functioning and conservation. Our results indicate diversity patterns to exhibit high variability, illustrating the complexity of driving factors. Nevertheless, we found that certain environmental variables can be defined as central for specific diversity metrics. The climate-dependency of the metrics suggests that different aspects of high-latitude diversity may face significant alterations due to projected climatic changes. However, our results highlight the significance of topo-edaphic variables in predicting diversity hotspots for species of elevated conservation concern, implicating certain conditions may sustain local populations of at-risk plants even under climate change.

Although the low hotspot congruence reduces proxy potential, understanding dissimilarities and parallels between diversity metrics could advance knowledge of the processes underlying diversity patterns and their relevance for conservation. Our findings have implications for future studies aiming to predict biodiversity as well as conservation decisions. Furthermore, we highlight the importance of employing comprehensive species and environmental data in predicting diversity and its manifestation across landscapes.

## Acknowledgements

AN was funded by the The Finnish Foundation for Nature Conservation, Societas pro Fauna et Flora Fennica, and The Doctoral Program in Geosciences. ML was partly funded by the Academy of Finland (Project Number 286950). We would also like to thank the Finnish Museum of Natural History as one of the main contributors of the data on vascular plant species, Juha Aalto for helping with gathering the environmental data, and Jianjun Wang for helping with the LCBD aspect of this work.

**Glossary**

Aster = advanced spaceborne thermal emission and reflection radiometer, an imaging instrument onboard NASA's Terra satellite

BRT = boosted regression tree model, combines the strengths of boosting and multiple regression trees

CR = critically endangered species according to the Red List, used here to build the THR variable

DEM = digital elevation model, here used to develop the topographic variables

Edaphic = relating to soil characteristics

EN = endangered species according to the Red List

FDD = freezing degree days, representing average overwintering conditions

GDD3 = growing degree days, representing average growing conditions

LCBD = local contribution to β-diversity of a site

NT = near threatened species according to the Red List, used here to build the THR variable

PA = protected area

RRR = (relative) range-rarity richness of a site

TSR = total species richness of a site

THR = richness of species of elevated conservation concern (based on the IUCN Red List) of a site

TSS = true skill statistic, a widely used measure of model performance

TWI = topographic wetness index

VU = vulnerable species according to the Red List, used here to build the THR variable

WAB = water balance, representing the range of available moisture

**Appendices**

Appendix A. The information on threat status was found from the Red List of Finnish Species (Rassi 2010) following the Red List Categories and Criteria (IUCN 2001) produced by the Species Survival Commission of the World Conservation Union (IUCN; http://www.iucn.org). The IUCN categories employed here are Critically Endangered (CR), Endangered (EN), Vulnerable (VU), and Near Threatened (NT). A total of 73 species (12%) were classified as threatened: three as CR, ten as EN, 23 as VU, and 37 as NT.

| **Vascular plant species name** | **IUCN category** | **No. of cells observed in** |
|---|---|---|
| *Alchemilla propinqua* | NT | 1 |
| *Allium schoenoprasum* | NT | 1 |
| *Antennaria dioica* | NT | 726 |
| *Antennaria nordhageniana* | VU | 18 |
| *Antennaria porsildii* | VU | 37 |
| *Antennaria villifera* | NT | 53 |
| *Arctagrostis latifolia* | NT | 2 |
| *Arenaria norvegica* | VU | 2 |
| *Armeria maritima* | EN | 1 |
| *Arnica angustifolia* | EN | 46 |
| *Botrychium boreale* | VU | 31 |
| *Botrychium lanceolatum* | VU | 6 |
| *Botrychium lunaria* | NT | 86 |
| *Botrychium multifidum* | NT | 16 |
| *Calypso bulbosa* | VU | 1 |
| *Campanula uniflora* | VU | 11 |
| *Carex appropinquata* | VU | 5 |
| *Carex atrata* | NT | 62 |
| *Carex atrofusca* | NT | 32 |
| *Carex disperma* | NT | 16 |
| *Carex fuliginosa* | NT | 25 |
| *Carex heleonastes* | VU | 19 |
| *Carex laxa* | NT | 37 |
| *Carex microglochin* | EN | 17 |
| *Carex rupestris* | NT | 69 |
| *Catabrosa aquatica* | NT | 1 |
| *Chamorchis alpina* | EN | 8 |
| *Cypripedium calceolus* | NT | 1 |
| *Dactylorhiza fuchsii* | NT | 18 |
| *Dactylorhiza lapponica* | VU | 4 |
| *Dactylorhiza traunsteineri* | VU | 7 |
| *Dianthus deltoides* | NT | 1 |
| *Dianthus superbus* | CR | 8 |
| *Draba alpina* | EN | 7 |
| *Draba daurica* | VU | 71 |
| *Draba fladnizensis* | VU | 14 |

| | | |
|---|---|---|
| *Draba lactea* | VU | 23 |
| *Draba nivalis* | NT | 45 |
| *Elymus fibrosus* | VU | 1 |
| *Epipogium aphyllum* | VU | 5 |
| *Erigeron borealis* | VU | 11 |
| *Erigeron humilis* | NT | 92 |
| *Eriophorum brachyantherum* | VU | 27 |
| *Euphrasia salisburgensis* | EN | 7 |
| *Gentianella tenella* | EN | 9 |
| *Hierochloë odorata* | NT | 15 |
| *Juncus arcticus* | EN | 17 |
| *Kobresia myosuroides* | CR | 2 |
| *Koenigia islandica* | NT | 82 |
| *Lappula deflexa* | VU | 4 |
| *Minuartia biflora* | NT | 179 |
| *Minuartia rubella* | VU | 16 |
| *Minuartia stricta* | VU | 20 |
| *Nardus stricta* | NT | 593 |
| *Oxytropis lapponica* | CR | 2 |
| *Pedicularis hirsuta* | NT | 148 |
| *Phippsia algida* | NT | 64 |
| *Polystichum lonchitis* | NT | 26 |
| *Potentilla chamissonis* | NT | 10 |
| *Potentilla nivea* | NT | 27 |
| *Pseudorchis albida* | NT | 29 |
| *Pyrola media* | NT | 5 |
| *Ranunculus glacialis* | NT | 213 |
| *Ranunculus sulphureus* | EN | 9 |
| *Rhododendron lapponicum* | NT | 41 |
| *Sagina nivalis* | NT | 73 |
| *Salix arbuscula* | EN | 2 |
| *Saxifraga hirculus* | VU | 23 |
| *Sedum villosum* | VU | 2 |
| *Silene wahlbergella* | NT | 46 |
| *Trisetum subalpestre* | NT | 28 |
| *Veronica fruticans* | NT | 36 |
| *Woodsia glabella* | NT | 20 |

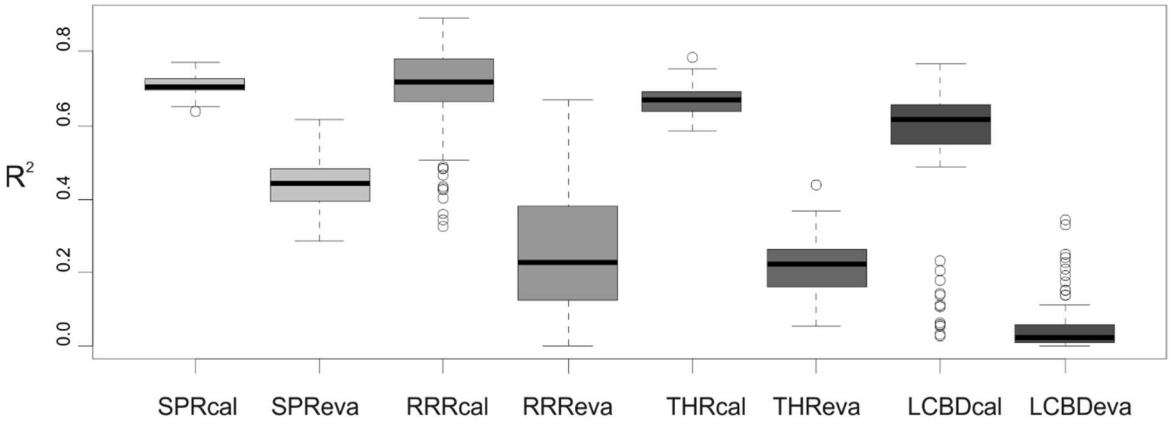
403

404

Appendix B. Bivariate Spearman correlation coefficient (ρ) estimates and corresponding p-values between the predictor variables. See Table 1 in main text for variable descriptions and abbreviations.

| | FDD | GDD3 | WAB | TWI | Slope | Rad | Calc | Evenness | Rock |
|---|---|---|---|---|---|---|---|---|---|
| FDD mean | ***** | -0.24 | 0.16 | 0.13 | 0.08 | 0.00 | 0.14 | 0.00 | 0.27 |
| GDD3 mean | <0.001 | ***** | -0.53 | -0.01 | -0.31 | -0.10 | -0.19 | 0.34 | -0.64 |
| WAB range | <0.001 | <0.001 | ***** | 0.22 | 0.70 | 0.30 | 0.13 | -0.08 | 0.65 |
| TWI range | <0.001 | 0.777 | <0.001 | ***** | 0.44 | 0.02 | 0.18 | -0.02 | 0.13 |
| Slope range | <0.001 | <0.001 | <0.001 | <0.001 | ***** | 0.22 | 0.21 | -0.03 | 0.47 |
| Rad mean | 0.967 | <0.001 | <0.001 | 0.296 | <0.001 | ***** | 0.03 | 0.00 | 0.10 |
| Calc mean | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | 0.256 | ***** | -0.19 | 0.13 |
| Soil div | 0.908 | <0.001 | 0.001 | 0.33 | 0.165 | 0.935 | <0.001 | ***** | -0.05 |
| Rock cover | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | 0.029 | ***** |

Appendix C. Boxplot showing model predictive performance, as demonstrated by $R^2$, for the calibration (cal) and evaluation (eva) data separately for each of the four diversity metrics. Model performance (tested with the Wilcoxon signed-rank test) was significantly ($p<0.001$) better for the training data for all metrics. See Table 1 and 2 in the main text for variable descriptions and abbreviations.



## References


Aalto, J. and M. Luoto. 2014. Integrating climate and local factors for geomorphological distribution models. Earth Surface Processes and Landforms 39:1729-1740.

Aalto, J., A. Venäläinen, R. K. Heikkinen, and M. Luoto. 2014. Potential for extreme loss in high-latitude Earth surface processes due to climate change. Geophysical Research Letters 41:3914-3924.

Ackerly, D. D., S. R. Loarie, W. K. Cornwell, S. B. Weiss, H. Hamilton, R. Branciforte, and N. J. B. Kraft. 2010. The geography of climate change: implications for conservation biogeography. Diversity and Distributions 16:476-487.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38:541-545.

Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43:1223-1232.

Anderson, M. G. and C. E. Ferree. 2010. Conserving the stage: climate change and the geophysical underpinnings of species diversity. PloS one 5:e11554.

Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J. Sanders, H. V. Cornell, L. S. Comita, and K. F. Davies. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology letters 14:19-28.

Armsworth, P. R., B. E. Kendall, and F. W. Davis. 2004. An introduction to biodiversity concepts for environmental economists. Resource and energy Economics 26:115-136.

Ashcroft, M. B. 2010. Identifying refugia from climate change. Journal of Biogeography 37:1407-1413.

Ashcroft, M. B. and J. R. Gollan. 2013. Moisture, thermal inertia, and the spatial distributions of near-surface soil and air temperatures: Understanding factors that promote microrefugia. Agricultural and Forest Meteorology 176:77-89.

Austin, M. P. and K. P. Van Niel. 2011a. Impact of landscape predictors on climate change modelling of species distributions: a case study with Eucalyptus fastigata in southern New South Wales, Australia. Journal of Biogeography 38:9-19.

Austin, M. P. and K. P. Van Niel. 2011b. Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography 38:1-8.

Barnosky, A. D. 2008. Climatic change, refugia, and biodiversity: where do we go from here? An editorial comment. Climatic Change 86:29-32.

Beven, K. J. and M. J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Bulletin 24:43-69.

Bomhard, B., D. M. Richardson, J. S. Donaldson, G. O. Hughes, G. F. Midgley, D. C. Raimondo, A. G. Rebelo, M. Rouget, and W. Thuiller. 2005. Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. Global Change Biology 11:1452-1468.

Bonn, A., A. S. Rodrigues, and K. J. Gaston. 2002. Threatened and endemic species: are they good indicators of patterns of biodiversity on a national scale? Ecology letters 5:733-741.

Brooks, T. M., R. A. Mittermeier, G. A. da Fonseca, J. Gerlach, M. Hoffmann, J. F. Lamoreux, C. G. Mittermeier, J. D. Pilgrim, and A. S. Rodrigues. 2006. Global biodiversity conservation priorities. Science 313:58-61.

Cañadas, E. M., G. Fenu, J. Peñas, J. Lorite, E. Mattana, and G. Bacchetta. 2014. Hotspots within hotspots: Endemic plant richness, environmental drivers, and implications for conservation. Biological Conservation 170:282-291.

Carter, T., J. Porter, and M. Parry. 1991. Climatic warming and crop potential in Europe: Prospects and uncertainties. Global Environmental Change 1:291-312.

Ceballos, G. and P. R. Ehrlich. 2006. Global mammal distributions, biodiversity hotspots, and conservation. Proceedings of the National Academy of Sciences 103:19374-19379.

464 Chape, S., J. Harrison, M. Spalding, and I. Lysenko. 2005. Measuring the extent and effectiveness of protected
465 areas as an indicator for meeting global biodiversity targets. Philosophical Transactions of the Royal
466 Society of London B: Biological Sciences 360:443-455.
467 Chytrý, M., L. Tichý, and J. Roleček. 2003. Local and regional patterns of species richness in Central European
468 vegetation types along the pH/calcium gradient. Folia Geobotanica 38:429-442.
469 Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychosocial Measurement,
470 20, 37-46.
471 Csuti, B., S. Polasky, P. H. Williams, R. L. Pressey, J. D. Camm, M. Kershaw, A. R. Kiester, B. Downs, R. Hamilton,
472 and M. Huso. 1997. A comparison of reserve selection algorithms using data on terrestrial
473 vertebrates in Oregon. Biological Conservation 80:83-97.
474 Daru, B. H., M. Bank, and T. J. Davies. 2015. Spatial incongruence among hotspots and complementary areas
475 of tree diversity in southern Africa. Diversity and Distributions 21:769-780.
476 Davies, R. G., C. D. L. Orme, D. Storch, V. A. Olson, G. H. Thomas, S. G. Ross, T.-S. Ding, P. C. Rasmussen, P. M.
477 Bennett, and I. P. Owens. 2007. Topography, energy and the global distribution of bird species
478 richness. Proceedings of the Royal Society of London B: Biological Sciences 274:1189-1197.
479 Dobrowski, S. Z. 2011. A climatic basis for microrefugia: the influence of terrain on climate. Global Change
480 Biology 17:1022-1035.
481 Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade,
482 and P. J. Leitão. 2013. Collinearity: a review of methods to deal with it and a simulation study
483 evaluating their performance. Ecography 36:27-46.
484 Dubuis, A., S. Giovanettina, L. Pellissier, J. Pottier, P. Vittoz, and A. Guisan. 2013. Improving the prediction of
485 plant species distribution and community composition by adding edaphic to topo-climatic variables.
486 Journal of Vegetation Science 24:593-606.
487 Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal
488 Ecology 77:802-813.
489 Esri. 2013. ArcScripts. www.arcsripts.esri.com
490 Feng, J.-M., Z. Zhang, and R.-Y. Nan. 2011. Non-congruence among hotspots based on three common
491 diversity measures in Yunnan, south-west China. Plant Ecology & Diversity 4:353-361.
492 Flather, C. H., K. R. Wilson, D. J. Dean, and W. C. McComb. 1997. Identifying gaps in conservation networks:
493 of indicators and uncertainty in geographic-based analyses. Ecological Applications 7:531-542.
494 Fronzek, S., M. Luoto, and T. R. Carter. 2006. Potential effect of climate change on the distribution of palsa
495 mires in subarctic Fennoscandia. Climate Research 32:1-12.
496 Fu, P. and P. M. Rich. 1999. Design and implementation of the Solar Analyst: an ArcView extension for
497 modeling solar radiation at landscape scales. Pages 1-31 in Proceedings of the Nineteenth Annual
498 ESRI User Conference.
499 Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405:220-227.
500 Gjerde, I., M. Sætersdal, J. Rolstad, H. H. Blom, and K. O. Storaunet. 2004. Fine-Scale Diversity and Rarity
501 Hotspots in Northern Forests. Conservation Biology 18:1032-1042.
502 Gould, W. 2000. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots.
503 Ecological Applications 10:1861-1870.
504 Guerin, G. R. and A. J. Lowe. 2015. 'Sum of inverse range-sizes'(SIR), a biodiversity metric with many names
505 and interpretations. Biodiversity and Conservation 24:2877-2882.
506 Guisan, A., J. P. Theurillat, and F. Kienast. 1998. Predicting the potential distribution of plant species in an
507 alpine environment. Journal of Vegetation Science 9:65-74.
508 Guisan, A. and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological
509 Modelling 135:147-186.
510 Harrison, S. and R. Noss. 2017. Endemism hotspots are linked to stable climatic refugia. Annals of
511 Botany:mcw248.
512 Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J.-F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T.
513 Oberdorff, and E. M. O'Brien. 2003. Energy, water, and broad-scale geographic patterns of species
514 richness. Ecology 84:3105-3117.

515 Heikkinen, R., H. Birks, and R. Kalliola. 1998. A numerical analysis of the mesoscale distribution patterns of
516 vascular plants in the subarctic Kevo Nature Reserve, northern Finland. Journal of Biogeography
517 25:123-146.
518 Heikkinen, R. K. and S. Neuvonen. 1997. Species richness of vascular plantsin the subarctic landscape of
519 northern Finland: modelling relationships to the environment. Biodiversity & Conservation 6:1181-
520 1201.
521 Huang, J., J. Huang, C. Liu, J. Zhang, X. Lu, and K. Ma. 2016. Diversity hotspots and conservation gaps for the
522 Chinese endemic seed flora. Biological Conservation 198:104-112.
523 Jetz, W. and C. Rahbek. 2002. Geographic range size and determinants of avian species richness. Science
524 297:1548-1551.
525 Kauhanen, H. O. 2013. Mountains of Kilpisjärvi host an abundance of threatened plants in Finnish Lapland.
526 Botanica Pacifica: A journal of plant science and conservation 2:43-52.
527 Kissling, W. D. and G. Carl. 2008. Spatial autocorrelation and the selection of simultaneous autoregressive
528 models. Global Ecology and Biogeography 17:59-71.
529 Lamoreux, J. F., J. C. Morrison, T. H. Ricketts, D. M. Olson, E. Dinerstein, M. W. McKnight, and H. H. Shugart.
530 2006. Global tests of biodiversity concordance and the importance of endemism. Nature 440:212-
531 214.
532 le Roux, P. C., J. Aalto, and M. Luoto. 2013a. Soil moisture's underestimated role in climate change impact
533 modelling in low-energy systems. Global Change Biology 19:2965-2975.
534 le Roux, P. C. and M. Luoto. 2014. Earth surface processes drive the richness, composition and occurrence of
535 plant species in an arctic–alpine environment. Journal of Vegetation Science 25:45-54.
536 le Roux, P. C., R. Virtanen, R. K. Heikkinen, and M. Luoto. 2012. Biotic interactions affect the elevational ranges
537 of high-latitude plant species. Ecography 35:1048-1056.
538 le Roux, P. C., R. Virtanen, and M. Luoto. 2013b. Geomorphological disturbance is necessary for predicting
539 fine-scale species distributions. Ecography 36:800-808.
540 Legendre, P. 2014. Interpreting the replacement and richness difference components of beta diversity. Global
541 Ecology and Biogeography 23:1324-1334.
542 Legendre, P. and M. Cáceres. 2013. Beta diversity as the variance of community data: dissimilarity coefficients
543 and partitioning. Ecology letters 16:951-963.
544 Lennon, J. J., P. Koleff, J. J. Greenwood, and K. J. Gaston. 2004. Contribution of rarity and commonness to
545 patterns of species richness. Ecology letters 7:81-87.
546 Lenoir, J., T. Hattab, and G. Pierre. 2016. Climatic microrefugia under anthropogenic climate change:
547 implications for species redistribution. Ecography.
548 Levin, N., A. Shmida, O. Levanoni, H. Tamari, and S. Kark. 2007. Predicting mountain plant richness and rarity
549 from space using satellite-derived vegetation indices. Diversity and Distributions 13:692-703.
550 Lobo, J. M., I. Castro, and J. C. Moreno. 2001. Spatial and environmental determinants of vascular plant
551 species richness distribution in the Iberian Peninsula and Balearic Islands. Biological Journal of the
552 Linnean Society 73:233-253.
553 Loidi, J., J. A. Campos, M. Herrera, I. Biurrun, I. García-Mijangos, and G. García-Baquero. 2015. Eco-
554 geographical factors affecting richness and phylogenetic diversity patterns of high-mountain flora in
555 the Iberian Peninsula. Alpine Botany 125:137-146.
556 Magurran, A. E. 2013. Ecological diversity and its measurement. Springer Science & Business Media.
557 Magurran, A. E. and M. Dornelas. 2010. Biological diversity in a changing world. Philosophical Transactions
558 of the Royal Society of London B: Biological Sciences 365:3593-3597.
559 Malcolm, J. R., C. Liu, R. P. Neilson, L. Hansen, and L. Hannah. 2006. Global warming and extinctions of
560 endemic species from biodiversity hotspots. Conservation Biology 20:538-548.
561 Maloufi, S., A. Catherine, D. Mouillot, C. Louvard, A. Couté, C. Bernard, and M. Troussellier. 2016.
562 Environmental heterogeneity among lakes promotes hyper β-diversity across phytoplankton
563 communities. Freshwater Biology 61:633-645.
564 Mawdsley, J. R., R. O'Malley, and D. S. Ojima. 2009. A review of climate-change adaptation strategies for
565 wildlife management and biodiversity conservation. Conservation Biology 23:1080-1089.

McCune, B. and D. Keon. 2002. Equations for potential annual direct incident radiation and heat load. Journal of Vegetation Science 13:603-606.
Médail, F. and K. Diadema. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. Journal of Biogeography 36:1333-1345.
Mod, H. K., D. Scherrer, M. Luoto, and A. Guisan. 2016. What we use is not what we know: environmental predictors in plant distribution models. Journal of Vegetation Science.
Moeslund, J. E., L. Arge, P. K. Bøcher, T. Dalgaard, and J.-C. Svenning. 2013. Topography as a driver of local terrestrial vascular plant diversity patterns. Nordic Journal of Botany 31:129-144.
Moritz, C. 2001. Biodiversity hotspots and beyond: the need for preserving environmental transitions. Trends in Ecology and Evolution 16.
Mossberg, B. and L. Stenberg. 2003. The new Nordic flora. Wahlstr om & Widstrand, Stockholm 928.
Mouillot, D., D. R. Bellwood, C. Baraloto, J. Chave, R. Galzin, M. Harmelin-Vivien, M. Kulbicki, S. Lavergne, S. Lavorel, N. Mouquet, C. E. Paine, J. Renaud, and W. Thuiller. 2013. Rare species support vulnerable functions in high-diversity ecosystems. PLoS Biol 11:e1001569.
Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. Da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853-858.
NASA Land Processes Distributed Active Archive Center. 2013. ASTER DEM.
Niskanen, A., M. Luoto, H. Väre, and R. K. Heikkinen. 2016a. Models of Arctic-alpine refugia highlight importance of climate and local topography. Polar Biology:1-14.
Niskanen, A. K., R. K. Heikkinen, H. K. Mod, H. Väre, and M. Luoto. 2016b. Improving forecasts of arctic-alpine refugia persistence with landscape-scale variables. Geografiska Annaler: Series A, Physical Geography:1-13.
Orme, C. D., R. G. Davies, M. Burgess, F. Eigenbrod, N. Pickup, V. A. Olson, A. J. Webster, T. S. Ding, P. C. Rasmussen, R. S. Ridgely, A. J. Stattersfield, P. M. Bennett, T. M. Blackburn, K. J. Gaston, and I. P. Owens. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436:1016-1019.
Pearson, R. G. and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global Ecology and Biogeography 12:361-371.
Penna, D., M. Borga, D. Norbiato, and G. Dalla Fontana. 2009. Hillslope scale soil moisture variability in a steep alpine terrain. Journal of Hydrology 364:311-327.
Peters, H., B. C. O'Leary, J. P. Hawkins, and C. M. Roberts. 2015. Identifying species at extinction risk using global models of anthropogenic impact. Global Change Biology 21:618-628.
Pirinen, P., H. Simola, J. Aalto, J. Kaukoranta, P. Karlsson, and R. Ruuhela. 2012. Climatological statistics of Finland 1981–2010. Finnish Meteorological Institute Reports 2012:25.
Possingham, H. P. and K. A. Wilson. 2005. Biodiversity: Turning up the heat on hotspots. Nature 436:919-920.
Prendergast, J., R. Quinn, J. Lawton, B. Eversham, and D. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365:335-337.
Randin, C. F., R. Engler, S. Normand, M. Zappa, N. E. Zimmermann, P. B. Pearman, P. Vittoz, W. Thuiller, and A. Guisan. 2009. Climate change and plant distribution: local models predict high-elevation persistence. Global Change Biology 15:1557-1569.
Rassi, P. 2010. Suomen lajien uhanalaisuus 2010: the 2010 red list of Finnish species. Ympäristöministeriö.
Reid, W. V. 1998. Biodiversity hotspots. Trends Ecol Evol 13:275-280.
Reside, A. E., J. A. Welbergen, B. L. Phillips, G. W. Wardell-Johnson, G. Keppel, S. Ferrier, S. E. Williams, and J. VanDerWal. 2014. Characteristics of climate change refugia for Australian biodiversity. Austral Ecology 39:887-897.
Ridgeway, G. 2010. gbm: generalized boosted regression models. R package version 1.6-3.1.
Scherrer, D. and C. Körner. 2011. Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. Journal of Biogeography 38:406-416.
Schut, A. G., G. W. Wardell-Johnson, C. J. Yates, G. Keppel, I. Baran, S. E. Franklin, S. D. Hopper, K. P. Van Niel, L. Mucina, and M. Byrne. 2014. Rapid characterisation of vegetation structure to predict refugia and climate change impacts across a global biodiversity hotspot. PloS one 9:e82778.

617 Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, and T. C.
618 Edwards Jr. 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife
619 monographs:3-41.
620 Skov, F. and J. C. Svenning. 2004. Potential impact of climatic change on the distribution of forest herbs in
621 Europe. Ecography 27:366-380.
622 Sormunen, H., R. Virtanen, and M. Luoto. 2011. Inclusion of local environmental conditions alters high-
623 latitude vegetation change predictions based on bioclimatic models. Polar Biology 34:883-897.
624 Steck, C. E., M. Bürgi, T. Coch, and P. Duelli. 2007. Hotspots and richness pattern of grasshopper species in
625 cultural landscapes. Biodiversity and Conservation 16:2075-2086.
626 Stein, B. A., L. S. Kutner, and J. S. Adams. 2000. Precious heritage: the status of biodiversity in the United
627 States. Oxford University Press on Demand.
628 Taberlet, P. and R. Cheddadi. 2002. Ecology. Quaternary refugia and persistence of biodiversity. Science
629 297:2009-2010.
630 Thuiller, W., S. Lavorel, M. B. Araújo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant
631 diversity in Europe. Proc Natl Acad Sci U S A 102:8245-8250.
632 UNEP-WCMC, I. a. 2016. The World Database on Protected Areas (WDPA) [On-line] [04/2016]. Cambridge,
633 UK: UNEP-WCMC.
634 Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon:213-251.
635 Williams, P. 2000. Some properties of rarity scores used in site quality assessment. British Journal of
636 Entomology and Natural History 13:73-86.
637 Williams, P., D. Gibbons, C. Margules, A. Rebelo, C. Humphries, and R. Pressey. 1996. A comparison of richness
638 hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds.
639 Conservation Biology 10:155-174.
640 Virkkala, R., R. K. Heikkinen, S. Fronzek, and N. Leikola. 2013. Climate change, northern birds of conservation
641 concern and matching the hotspots of habitat suitability with the reserve network. PloS one
642 8:e63376.
643 Wohlgemuth, T. 1998. Modelling floristic species richness on a regional scale: a case study in Switzerland.
644 Biodiversity & Conservation 7:159-177.
645 Zellweger, F., V. Braunisch, F. Morsdorf, A. Baltensweiler, M. Abegg, T. Roth, H. Bugmann, and K. Bollmann.
646 2015. Disentangling the effects of climate, topography, soil and vegetation on stand-scale species
647 richness in temperate forests. Forest Ecology and Management 349:36-44.

648

649

650 **Tables**

651 Table 1. Descriptions of the four indices used in this study for estimating biodiversity

| Diversity index | Abbreviation & Equation | Details | Characteristics | Reference |
|---|---|---|---|---|
| Total species richness | $\mathrm{TSR} = S$ | S is the number of vascular plant species within a grid cell | Straightforward and universal | Whittaker 1972 |
| Relative range-rarity richness | $RRR = \frac{\sum_{i=1}^{n} Wi}{\mathrm{TSR}}$ | *n* is the number of species in a grid cell, *Wi* is the weighting of species *i*, here the inverse of its range, TSR as above | Combines richness with range size, accounts for total species richness | Williams et al. 1996 |
| Threatened/near-threatened species richness | $THR = THS$ | *THS* is the number of threatened or near-threatened vascular plant species within a grid cell | Combines threat risk of Red Listed species | Gjerde et al. 2004 |
| Local contribution to β-diversity | $LCBD = \frac{SS_i}{SS_{Total}}$ | $SS_i$ is the sum of squares of the *i*th sampling unit, $SS_{total}$ is the sum of squares of the species data | Gives the relative contribution of a site to β-diversity | Legendre et al. 2013 |

652

653

654 Table 2. Descriptions of the variables (response variables above dashed line, environmental variables below)
655 with their minimum (Min), median (Med), mean, and maximum (Max) values

| **Variable** | **Abbreviation** | **Min** | **Med** | **Mean** | **Max** |
|---|---|---|---|---|---|
| Total species richness | TSR | 6 | 90 | 90 | 240 |
| Relative range-rarity richness | RRR | 0 | 0.03 | 0.34 | 6.24 |
| Threatened/near-threatened species richness | THR | 0 | 1 | 1.34 | 32 |
| Local contribution to β-diversity | LCBD | 0.0003 | 0.0005 | 0.0005 | 0.0007 |
| Freezing degree days (annual accumulated daily temperature sum <0°C; unit °C) | FDD | -2062 | -1776 | -1769 | -1514 |
| Growing degree days (annual accumulated daily temperature sum >3°C; unit °C) | GDD3 | 109.3 | 553.8 | 556 | 1118 |
| Water balance (mm) | WAB | 2.2 | 9.5 | 11.8 | 60.8 |
| Range of slope angle | Slope | 11.2 | 25.5 | 27.4 | 77.5 |
| Potential annual direct radiation (MJ/cm²/a) | Rad | 0.3 | 0.4 | 0.4 | 0.6 |
| Topographic wetness index | TWI | 0.9 | 1.8 | 1.9 | 6.3 |
| Cover of calcareous substrates (%) | Calc | 0 | 9.6 | 14.6 | 90.3 |
| Cover of rocky substrate (%) | Rock | 0 | 0.1 | 0.2 | 1 |
| Evenness of substrate type (Simpson's *E*) | Soil | 0.3 | 0.4 | 0.4 | 0.9 |

656

657

## Figures

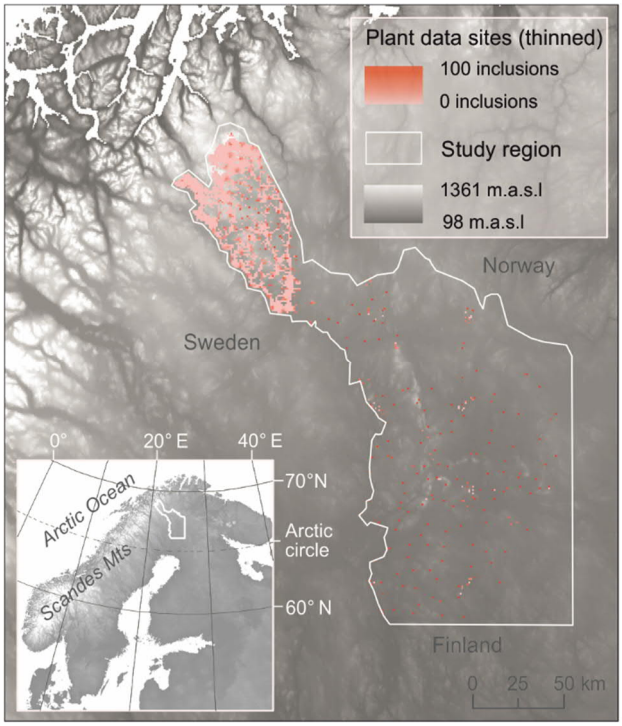



Figure 1. The location and elevation of the study area in northernmost Europe. The grid cells with available vascular plant data ($n$ =2073, spatial resolution 1 × 1 km) for which data thinning was applied are also shown. In total, 812 data points were included in the models.

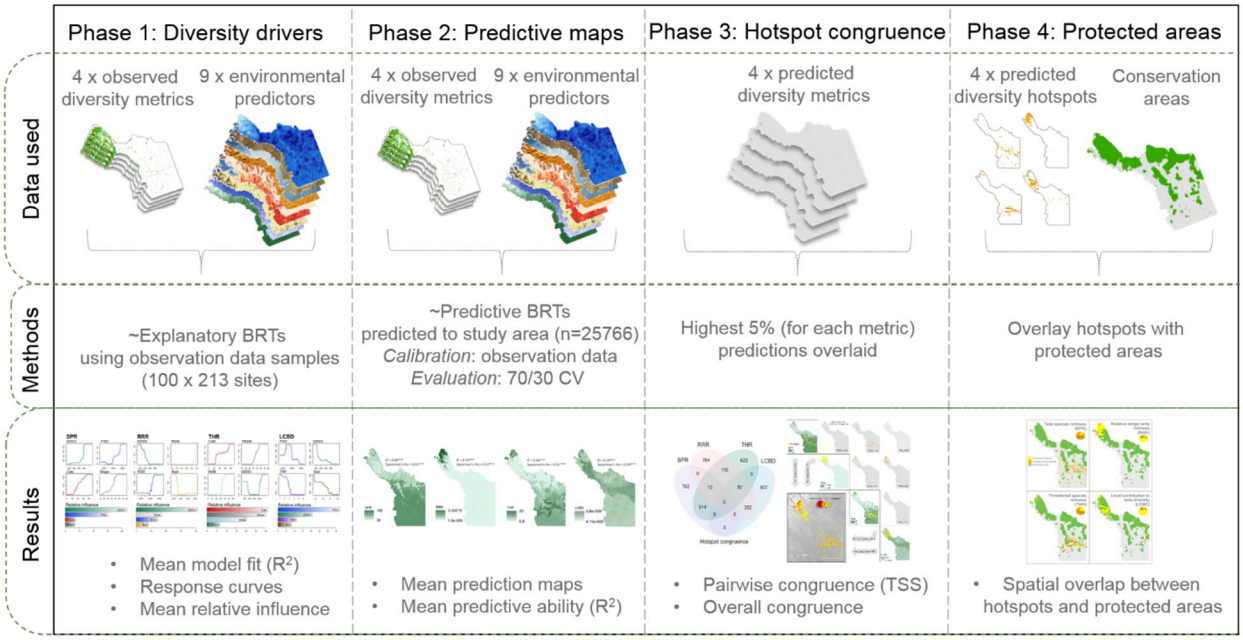
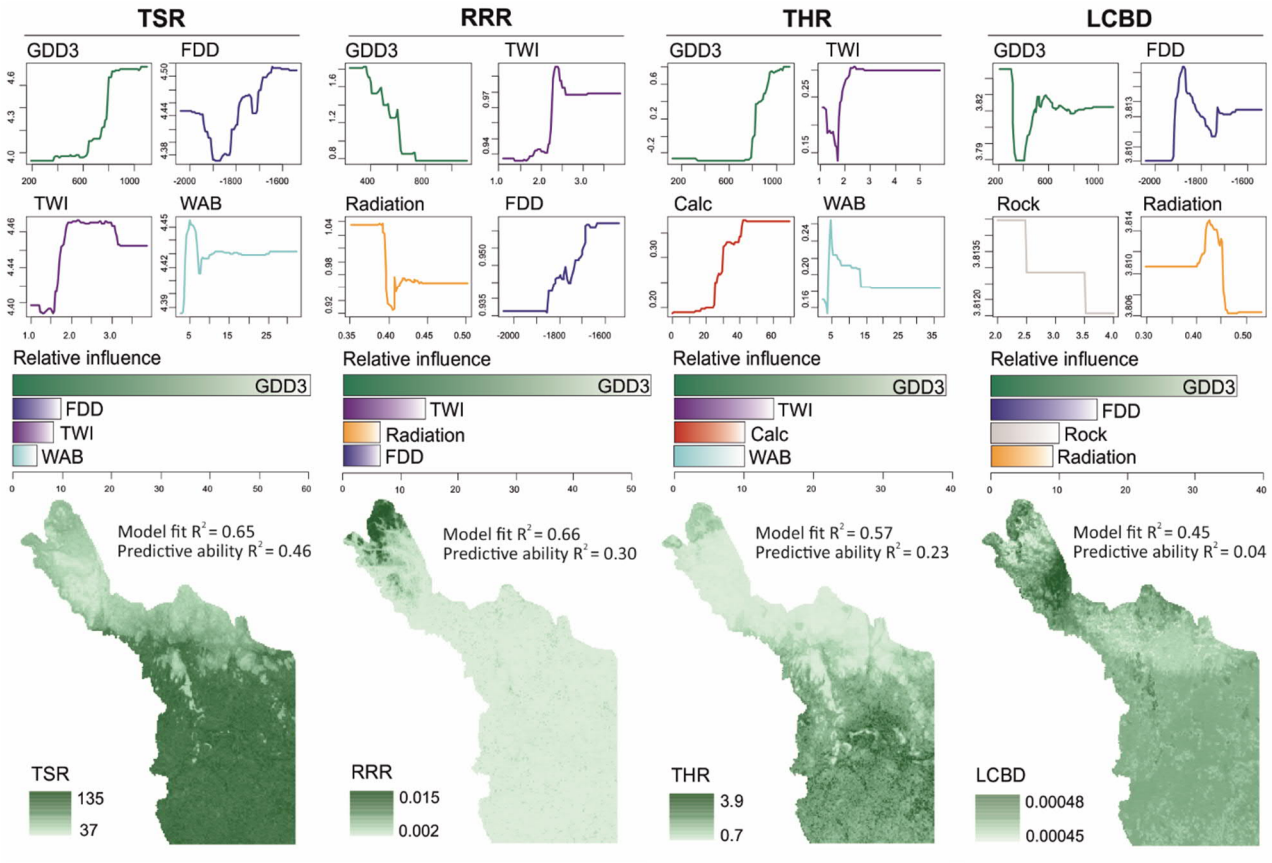
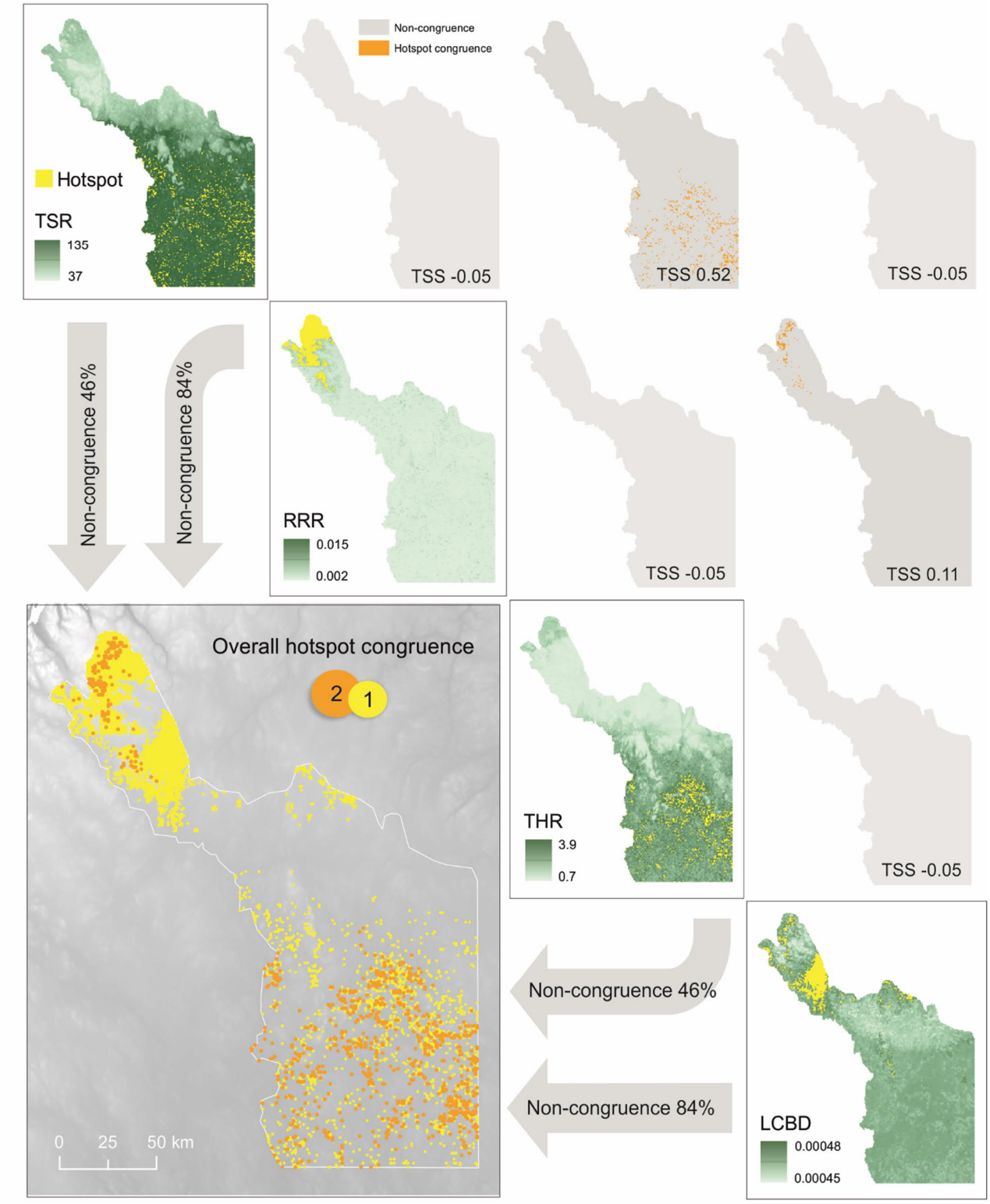



Figure 2. An overview of the data and methods used in each of the four phases of the modelling framework employed here, from building the boosted regression trees (BRT) to overlaying predictions with existing protected areas, and their results




668

669 Figure 3. Boosted regression tree (BRT) based results and predicted diversity maps for total species richness
670 (TSR), range-rarity richness (RRR), threatened species richness (THR), and local contribution to β-diversity
671 (LCBD). For each metric, a five-panel plot shows the BRT results: the four topmost panels are partial
672 dependency plots showing the BRT-modelled responses of the metric to its most influential drivers according
673 to the model with the highest explanatory power ($R^2$); the Relative influence panel shows the mean relative
674 contributions (%) of the four most influential variables in predicting richness values. The map below the plot
675 panels shows mean model predictions for each metric separately. See Table 2 for variable descriptions and
676 abbreviations




677

678 Figure 4. Model results and hotspot congruence. The distribution patterns of the four diversity metrics and
679 their hotspots are shown in the middle panels; pairwise congruence is shown with corresponding TSS values
680 in the upper panels; and, arrows, showing the portion of non-congruence of individual metrics with other
681 metrics, point to a map of all hotspots and their congruence (lower left panel). See Table 1 for variable
682 descriptions and Table 2 for abbreviations

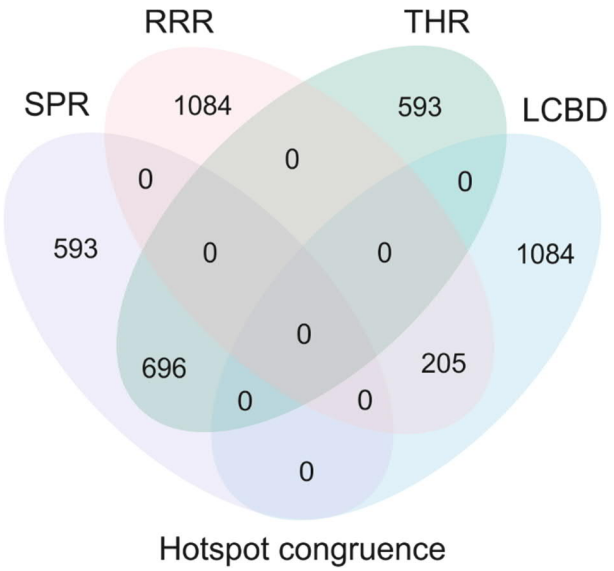
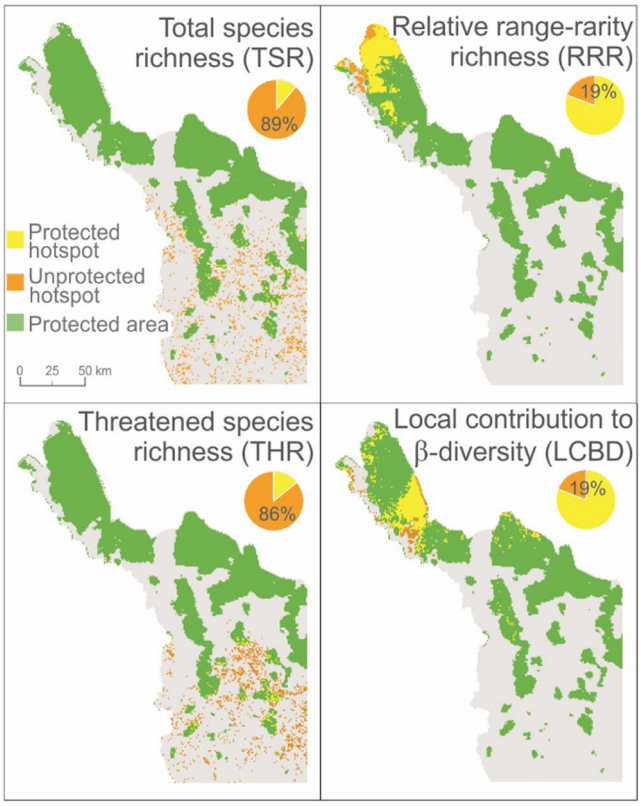



Figure 5. Spatial overlap and mismatch for the different hotspot components. Hotspots occupy a total of 4255 grid cells of which 3354 are hotspots for one metric only. Overall hotspot congruence is thus low: 901 hotspots were congruent for two metrics, with no congruent found for three or more metrics. See Table 1 for variable descriptions and Table 2 for abbreviations




688

689 Figure 6. Maps showing the congruence (in yellow) between the extent of the current protected area network
690 in the study region and the predicted diversity hotspots. The non-congruence (in orange), i.e. hotspots not
691 protected by protected areas, is also expressed in the inserted charts